Performance of methods for estimating size–transition matrices using tag–recapture data

André E. Punt[A,B,F], *Rik C. Buckworth*[C], *Catherine M. Dichmont*[D]
and Yimin Ye[D,E]

[A]CSIRO Marine and Atmospheric Research, GPO Box 1538, Hobart, Tas. 7001, Australia.

[B]School of Aquatic and Fishery Sciences, Box 355020, University of Washington,
Seattle, WA 98195-5020, USA.

[C]Fisheries, Department of Regional Development, Fisheries and Resources, GPO Box 3000,
Darwin, NT 0810, Australia.

[D]CSIRO Marine and Atmospheric Research, PO Box 120, Cleveland, Qld 4163, Australia.

[E]Current address: Fishery Management & Conservation Service, FAO of the United Nations,
Viale delle Termi diCaracalla, 00153, Rome, Italy.

[F]Corresponding author. Email: andre.punt@csiro.au

Abstract. Management advice for hard-to-age species such as prawns, crabs and rock lobsters are usually based on size-structured population dynamics models. These models require a size–transition matrix that specifies the probabilities of growing from one size-class to the others. Many methods exist to estimate size–transition matrices using tag–recapture data. However, they have not been compared in a systematic way. Eight of these methods are compared using Monte Carlo simulations parameterised using the data for the tiger prawn (*Penaeus semisulcatus*). Four of the methods are then applied to tag–recapture data for three prawn species in Australia's Northern Prawn Fishery to highlight the considerable sensitivity of model outputs to the method for estimating the size–transition matrix. The simulations show that not all methods perform equally well and that some methods are extremely poor. The 'best' methods, as identified in the simulations, are those that allow for individual variability in the parameters of the growth curve as well as the age-at-release. A method that assumes that ℓ_∞ rather than k varies among individuals tends to be more robust to violations of model assumptions.

Additional keywords: Australia, prawns, size-structured models, tagging data.

Introduction

Crustacean species support some of the most commercially valuable fisheries worldwide. For example, there are major fisheries for crabs, lobsters and prawns in many regions of the world and FAO fishery statistics indicate that almost 6000 kt of crustaceans were caught during 2006, 7% of global marine production (Food and Agriculture Organization 2008). Within Australia, crustaceans form the basis for the most valuable State (Western Australian rock lobster, de Lestang and Melville-Smith 2006) and Commonwealth (prawn) fisheries (Galeano *et al*. 2006; McLoughlin 2006). A wide variety of methods of stock assessment has been used to provide scientific advice for crustacean fisheries (Smith and Addison 2003). Of these methods, the most sophisticated attempt to track the size of stocks over time rather than simply assessing equilibrium relationships between yield and fishing mortality. However, stock assessments for many crustacean species cannot be based on age-structured population dynamics models and must be based instead on size-structured models because of difficulties assigning ages to individual animals (e.g. Hobday and Punt 2001; Kim *et al*. 2004; Chen *et al*. 2005). A key feature of size-structured stock assessment methods is the size–transition matrix, i.e. the matrix that defines the

probability of an animal growing from one size-class to the others over a certain time-step. This matrix is analogous to the growth curve in conventional age-structured stock assessment methods, although it plays a more fundamental role in determining the dynamics of the population in size-structured stock assessment methods.

Considerable focus, including extensive simulation studies, has been placed on estimating the parameters of growth curves (usually, but not always, the von Bertalanffy growth curve), specifically in the presence of individual variation in growth (e.g. Sainsbury 1980; Wang *et al*. 1995; Troynikov 1998; Laslett *et al*. 2002; Eveson *et al*. 2007). However, this previous work has concentrated on estimating the parameters determining the mean length-at-age of individuals in the population rather than the information needed to apply size-structured stock assessment methods, i.e. the entries of the size–transition matrix. These entries are defined not only in terms of the mean growth increment as a function of size, but also in terms of the variability associated with this increment.

A variety of methods has been proposed for estimating size–transition matrices for use in size-structured stock assessments (e.g. Sullivan *et al*. 1990; Punt *et al*. 1997, 2006; McGarvey *et al*.

1999; McGarvey and Feenstra 2001; Breen *et al*. 2003; Brandão *et al*. 2004). All of these methods are based on fitting growth models to tag–recapture data. They differ in terms of the form of the underlying growth curve, how individual variation in the growth increment is treated and the method of parameter estimation. For example, the methods of McGarvey *et al*. (1999), McGarvey and Feenstra (2001), Brandão *et al*. (2004) and Breen *et al*. (2003) defined the individual variability about the mean growth increment as additive error (either with a constant standard deviation or a constant coefficient of variation), whereas Punt *et al*. (2006) assumed that the von Bertalanffy *k* parameter varies among individuals according to a pre-specified probability distribution (either gamma, log-normal or truncated normal). Punt *et al*. (1997) assumed additive error within one time-step, but that the distribution of growth increment for *t* time-steps at liberty is the *t*th power of the size–transition matrix, thereby directly mimicking the structure of the size-structured assessment model in which the size–transition matrix is to be used. Several authors (e.g. Breen *et al*. 2003) have attempted to integrate the estimation of the size–transition matrix directly within the stock assessment model.

The present paper uses Monte Carlo simulations based on female *Penaeus semisulcatus* De Haan in Australia's Northern Prawn Fishery to compare eight methods grouped within three classes for estimating size–transition matrices. It then uses the best methods within each class to estimate size–transition matrices for three of the prawn species found in this fishery (*P. semisulcatus*, *P. esculentus* Haswell and *Metapenaeus endeavouri* Schmitt), highlighting diagnostic statistics that should be considered when fitting models to estimate size–transition matrices. All lengths referred to in this paper are carapace lengths (CL, mm), but the results are insensitive to the choice of the length measurement.

Materials and methods

Alternative methods for estimating the size transition matrix

The size transition matrix, $\mathbf{X}$, defines the probability, $X_{i,j}$, of growing from one size-class, i, to each of the size-classes in the model, $j = 1, 2, \ldots, m$, where m is the number of size-classes in the model, over a time-step Δt. The values for the parameters that determine the size–transition matrix are based on fitting models to tag–recapture data. The probability of an animal of size l_p^r growing to size l_p^c over a time period Δt_p is based on the assumption that the growth of an individual follows a von Bertalanffy growth curve:

$$l_p^c = l_p^r + (\ell_{\infty,p} - l_p^r)(1 - e^{-k_p \Delta t_p}) \qquad (1)$$

where l_p^r is the size of the *p*th animal when it was released, l_p^c is the size of the *p*th animal when it was recaptured, Δt_p is the time for which the *p*th animal was at liberty, k_p is the von Bertalanffy growth rate parameter for the *p*th animal and $\ell_{\infty,p}$ is the asymptotic size for the *p*th animal.

The estimable parameters of Eqn 1 are k_p and $\ell_{\infty,p}$. Although the results of the current paper are based on the von Bertalanffy growth curve, the methods can be applied straightforwardly

to other growth curves. All of the methods outlined below are recaptured–conditioned, i.e. animals that were released, but not subsequently recaptured, are ignored when estimating parameters.

Class 1 – k or ℓ_∞ varies among individuals

This class of method is based on the assumption that one of the parameters of the growth curve (*k* or ℓ_∞ for the von Bertalanffy growth curve) varies among individuals according to some pre-specified probability distribution (Wang *et al*. 1995; Troynikov 1998), while the value of the other parameter is the same for all animals, i.e.:

$$k_p \sim g(\cdot) \text{ or } \ell_{\infty,p} \sim h(\cdot) \qquad (2)$$

where g is the probability distribution for the von Bertalanffy growth rate parameter and h is the probability distribution for ℓ_∞.

The likelihood, L, in Eqn 3*a* is maximised to estimate the value of ℓ_∞ for all individuals and the parameters of g when k is assumed to differ among individuals, while L in Eqn 3*b* is maximised to estimate the value of k for all individuals and the parameters of h when ℓ_∞ is assumed to differ among individuals:

$$L = \prod_p \frac{g(\tilde{k}_p)}{\Delta t_p (\ell_\infty - l_p^c)}. \qquad (3a)$$

$$L = \prod_p \frac{h(\tilde{\ell}_{\infty,p})}{1 - e^{-k \Delta t_p}} \qquad (3b)$$

$$\tilde{k}_p = -\ell n[1 - (l_p^c - l_p^r)/(\ell_\infty - l_p^r)]/\Delta t_p;$$
$$\tilde{\ell}_{\infty,p} = [l_p^c - l_p^r e^{-k\Delta t_p}]/[1 - e^{-k\Delta t_p}]$$

where $\tilde{k}_p$ and $\tilde{\ell}_{\infty,p}$ are, respectively, the values of *k* for each individual (Eqn 3*a*) and ℓ_∞ for each individual (Eqn 3*b*) given the sizes at release and recapture and the estimates of the population mean values for ℓ_∞ and *k* respectively. The denominators of Eqns 3*a* and 3*b* are Jacobian terms that represent the transformation of l^c to *k* (i.e. $|\frac{dk}{dl^c}|$) and l^c to ℓ_∞ (i.e. $|\frac{d\ell_\infty}{dl^c}|$), respectively.

Wang *et al*. (1995) noted that Eqns 3*a* and 3*b* fail to account for the fact that the age-at-release is a random variable, that the variance of the growth increment depends on the age-at-release and that estimates of the parameters of the growth curve will be biased if account is not taken of this source of variation. They developed a likelihood function that assumes that the age-at-release, a_p, is distributed according to a distribution *f*. The joint likelihood functions for l_p^c and l_p^l for the cases in which *k* or ℓ_∞ are assumed to vary among individuals and a_p is treated as a random variable are given by:

$$L = \prod_p \frac{g(\tilde{k}_p) f(\tilde{a}_p)}{\Delta t_p (\ell_\infty - l_p^c) \ell_\infty \tilde{k}_p (1 - l_p^r/\ell_\infty)} \qquad (3c)$$

$$L = \prod_p \frac{h(\tilde{\ell}_{\infty,p}) f(\tilde{a}_p)}{l_p^c - l_p^r e^{-k\Delta t_p} k(1 - l_p^r/\tilde{\ell}_{\infty,p})} \quad (3d)$$

$$a_p = -\frac{1}{k_p} \ell n(1 - l_p^r/\ell_{\infty,p})$$

As for Eqns 3*a* and 3*b*, the numerators are the evaluations of *g* and *h* for $\tilde{k}_p$ and $\tilde{\ell}_{\infty,p}$, except that in this case, these terms are multiplied by an evaluation of *f* for $\tilde{a}_p$. The denominators of Eqns 3*c* and 3*d* are the Jacobians of the transformations between the lengths at release and recapture and the random variables included in *f*, *g* and *h*. The distributions *f*, *g* and *h* are taken to be two-parameter gamma functions for the purposes of this paper, although the basic approach can, in principle at least, be applied assuming *f*, *g* and *h* are any probability distribution based on the positive real numbers. The gamma distribution was chosen for this study because it has been used in several previous studies (e.g. Troynikov 1998).

The values for the entries of the size–transition matrix are evaluated using Monte Carlo integration for the methods in this class because, unlike the case for classes #2 and #3 below, calculation of the size–transition matrix cannot be achieved analytically. For methods based on Eqns 3*a* and 3*b*, and a given size-class *i*, this involves (1) generating sizes *L* for a large number of animals in size-class *i*, (2) computing the size of each of these animals after a time Δt based on values for k / ℓ_∞ simulated from their estimated probability distributions using the expression $(\ell_\infty - L)(1 - e^{-k\Delta t})$, and (3) computing the proportion of the animals that grew into each size-class *j*. In contrast, for methods based on Eqns 3*c* and 3*d*, the entries for size-class *i* are calculated by first generating values from *g*, *h* and *f*, excluding all animals not in size-class *i*, and applying steps (2) and (3) above.

Class 2 – fitting the size–transition matrix directly

After grouping the data to the width of each size-class and the time-step in the model that will be used for stock assessment purposes, the tag–recapture data can be summarised by sets of triplets (l_1, *t* and l_2, where l_1 is the length-at-release, *t* is the time-at-liberty and l_2 is the length-at-recapture). The likelihood function is then the product over animals of the probability of observing that a prawn tagged at length l_1 and at liberty for *t* time-steps was recaptured at length l_2 (Punt *et al.* 2006). This probability is the (l_1, l_2) entry of the matrix $\mathbf{X}^t$. For the purposes of the present study, it is assumed that the probability that an animal in size-class *i* grows into size-class *j* during each time-step is governed by a normal distribution in which the 'left' and 'right' variance parameters are different, i.e.

$$X_{i,j} = \Omega_{i,j} \bigg/ \sum_{j'} \Omega_{i,j'}$$

$$\Omega_{i,j} = \begin{cases} \exp\left(-\frac{\{L_j - (L_i + I_i)\}^2}{2\sigma_1^2}\right) & \text{if } L_j - L_i > I_i \\ \exp\left(-\frac{\{L_j - (L_i + I_i)\}^2}{2\sigma_2^2}\right) & \text{otherwise} \end{cases} \quad (4)$$

where σ_1 and σ_2 determine the variability in the growth increment, L_i is the midpoint of size-class *i*, L_j is the midpoint of size-class *j* and I_i is the growth increment for an animal in size-class *i* according to a von Bertalanffy growth curve parameterised in terms of the modal values for *k* and ℓ_∞.

Any probability distribution can form the basis for the matrix Ω in Eqn 4; this particular 'normal' distribution was chosen because it is more flexible than many other candidate distributions (such as the gamma and log-normal distributions). In principle, alternative distributions could be examined and model selection criteria such as Akaike's Information Criterion (AIC: Akaike 1974; Burnham and Anderson 2002) or the Bayesian Information Criterion (BIC: Schwarz 1978) used to select among them.

Class 3 – specified standard deviation method

Maunder (2001) and Breen *et al.* (2003) estimated the size–transition matrix under the assumption that the expected growth increment for an animal released at a size of l_p^r is $I_p = (\ell_\infty - l_p^r)(1 - e^{-k\Delta t_p})$ and that the standard deviation of the growth increment, $l_p^r - l_p^c$, is either:

(1) linearly proportional to the expected growth increment, i.e. $\sigma_p = \phi I_p$;
(2) a power function of the expected growth increment, i.e. $\sigma_p = \phi(I_p)^\psi$; or
(3) the square root of the sum of a constant term and a term that is linearly proportional to the square of the expected growth increment, i.e. $\sigma_p = \sqrt{(\phi I_p)^2 + \lambda^2}$.

Given a specification for the standard deviation of the growth increment, the values for the parameters of the growth transition matrix are estimated by maximising the following likelihood function:

$$L = \prod_p \frac{1}{\sqrt{2\pi}\sigma_p} e^{-\frac{(l_p^c - (l_p^r + I_p))^2}{2(\sigma_p)^2}} \quad (5)$$

while the elements of the size–transition matrix for a time-step of length Δt are calculated using the equation:

$$X_{i,j} = \int_0^h \frac{1}{\sqrt{2\pi}\sigma(I_i)} \exp\left[-\frac{\{L_j + l - (L_i + I_i)\}^2}{2(\sigma(I_i))^2}\right] dl \quad (6)$$

where *h* is the width of a size-class in the assessment model (1 cm for the purposes of the analyses of this paper), L_k is the midpoint of size-class *k* (for size-classes $k = i$ and *j*) and I_i is the growth increment for the *i*th size-class based on Eqn 1.

Simulation evaluation

The three classes of methods lead to a total of eight potential estimation methods (Table 1). An evaluation of these eight methods is based on generating pseudo tag–recapture data from a simulated prawn population, which is designed to represent the attributes of female *P. semisulcatus* in Australia's Northern Prawn Fishery as it is understood based on the literature. This population is assumed to have a rate of total mortality of 0.045 week^{-1} (set, for simplicity, to the value assumed for

Table 1. Definition of the three classes and eight methods (A–H) considered in this paper

Class 1: k or ℓ_∞ varies among individuals
(A) individual variation in k (Eqn 3*a*)
(B) individual variation in ℓ_∞ (Eqn 3*b*)
(C) individual variation in k and a (Eqn 3*c*)
(D) individual variation in ℓ_∞ and a (Eqn 3*d*)

Class 2: Fitting the size transition matrix directly
(E) Eqn 4

Class 3: Specified standard deviation method
(F) $\sigma_p = \phi I_p$
(G) $\sigma_p = \phi (I_p)^\psi$
(H) $\sigma_p = \sqrt{(\phi I_p)^2 + \lambda^2}$

**Table 2. The parameters that define the simulated population**

Scenario #	k	ℓ_∞
1	$\bar{k} = 0.032\,\text{week}^{-1}$; $CV_k = 0.6$	$\ell_\infty = 55\,\text{mm}$
2	$k = 0.06\,\text{week}^{-1}$	$\bar{\ell}_\infty = 48\,\text{mm}$; $CV_{\ell_\infty} = 0.1$
3	$\bar{k} = 0.032\,\text{week}^{-1}$; $CV_k = 0.6$	$\bar{\ell}_\infty = 48\,\text{mm}$; $CV_{\ell_\infty} = 0.1$

the rate of natural mortality in assessments of tiger prawns in Australia's Northern Prawn Fishery, e.g. Dichmont *et al*. 2003). Three scenarios are considered: ℓ_∞ differs among individuals (scenario #1), k differs among individuals (scenario #2) and both ℓ_∞ and k differ among individuals (scenario #3). A base-case set of simulations was conducted along with a set of sensitivity tests.

The specifications of the simulations

The individual variability in ℓ_∞ and k was assumed to be gamma-distributed for the base-case simulations (see Table 2 for the parameters assumed for the various scenarios). The values in Table 2 for scenarios #1 and #2 were selected based on the application of methods C and D in Table 1 to data for *P. semisulcatus* females (the results based on methods A and B were qualitatively identical). The parameter choices for scenario #3 lead to more individual variation in growth than for scenarios #1 and #2, and scenario #3 consequently reflects a 'more difficult' case.

The prawns in the simulated population were assumed to have been spawned randomly with respect to week given the implications of the assumed rate of total mortality for the base-case set of simulations. The ages of the prawns in this population at the time of tagging were therefore generated as a sample from an exponential distribution. Each animal was assigned values for its growth parameters based on the specifications for the scenario being examined. The tag–recapture data were then generated subject to the simulated length-at-release being greater than 25 mm (roughly the smallest female *P. semisulcatus* that is typically captured in the fishery) and a logistic selection pattern (for both release and recapture) with a length-at-50%-selectivity of 40 mm, and 5th and 95th percentiles of 30.2 and 49.8 mm (based on the fraction of tagged prawns of different sizes released and later recaptured). The time-at-liberty for each released prawn was drawn from the actual times-at-liberty for *P. semisulcatus*
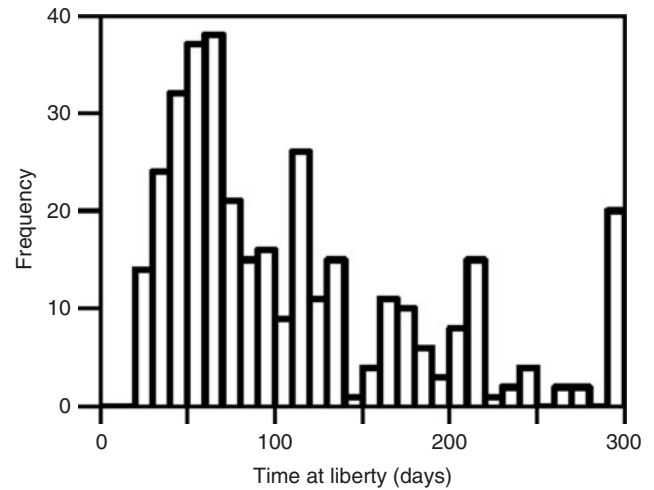
Fig. 1. Frequency distribution of the time-at-liberty (restricted to two weeks or more) for tagged female *Penaeus semisulcatus* in Australia's Northern Prawn Fishery.

females in Australia's Northern Prawn Fishery, restricted to animals that were at large for at least 14 days to reflect the standard restriction to allow tagged animals to fully mix into the population (Fig. 1). No measurement error was assumed for the base-case simulations, i.e. tagged and recaptured prawns were assumed to be measured perfectly. The base-case simulations were based on a sample size of 350 recaptures (the number of data points actually available for *P. semisulcatus* females).

The behaviour of the eight methods was further examined by changing some of the specifications of the base-case simulations (see Table 3 for details), in addition to changing the sample size from 350 to 50, 100, 200 and 600.

Evaluation of performance

The performance of the eight methods was evaluated in terms of the ability to estimate (a) the population mean (among individuals) values for ℓ_∞ and k, (b) the mean size of a 25-mm prawn after 10, 20, 30, etc. weeks, (c) the entries of the size–transition matrix, and (d) the equilibrium size distribution of a prawn population for which the rate of total mortality is $0.045\,\text{week}^{-1}$ and for which all recruitment is to the first size-class. The equilibrium size-distribution is the right eigenvector of the product of the size–transition matrix and a matrix with $\exp(-0.045)$ on the diagonal. Although the ability to estimate the population mean (between individuals) values for ℓ_∞ and k was of general interest, the results and discussion focus on quantities (b)–(d) because they relate most directly to applying size-structured stock assessment methods.

Estimation performance for the population mean values for ℓ_∞ and k was summarised by the distribution for the relative error, i.e. for simulation *v*:

$$100 \left(\frac{Y^{True} - Y_v^{Est}}{Y^{True}} \right) \qquad (7)$$

where Y^{True} is the true value of the quantity being estimated and Y_v^{Est} is the estimate of Y^{True} for simulation *v*.

Table 3.   The sensitivity tests and their abbreviations

Abbreviation	Description
Obs	Prawns are measured subject to normally distributed measurement error with a standard deviation of 1 mm
Prc-1	Prawn growth is stochastic* with a coefficient of variation, α, of 0.05
Prc-2	Prawn growth is stochastic* with a coefficient of variation, α, of 0.1
Logn	ℓ_∞/k are log-normally distributed rather than being gamma distributed
CoCV	ℓ_∞/k are normally distributed (with a constant coefficient of variation) rather than being gamma distributed
All20	All of the tagged prawns were spawned 20 weeks before the start of the tagging experiment
No Sel	Selectivity of the harvest is independent of prawn size

*Implemented by calculating growth increments for each week for which the prawn was at liberty and adding them together; the weekly growth increment is normally distributed with mean $(\ell_\infty - L)(1 - e^{-k})$ and standard deviation $\alpha\ell_\infty$.

Estimation performance for the mean size of a 25-mm animal after $10, 20, \ldots, 70$ weeks was quantified by the average absolute relative error over weeks:

$$\frac{100}{7} \sum_j \frac{|\overline{L}_j^{True} - \overline{L}_{j,v}^{Est}|}{\overline{L}_j^{True}} \tag{8}$$

where $\overline{L}_j^{True}$ is the true mean size of a 25-mm animal after $10j$ weeks, and $\overline{L}_{j,v}^{Est}$ is the estimate of $\overline{L}_j^{True}$ for simulation v. Results were obtained for 25-, 30-, 35-, 40- and 45-mm animals, but are only reported for 25-mm animals because the results for the other sizes were qualitatively identical to those for 25-mm animals.

Estimation performance for the entries of the size–transition matrix was quantified by the sum of the absolute errors, i.e. for simulation v:

$$\frac{100}{2m} \sum_i \sum_j |X_{i,j}^{True} - X_{i,j,v}^{Est}| \tag{9}$$

where $X_{i,j}^{True}$ is the true value of the (i,j) element of the size–transition matrix, $X_{i,j,v}^{Est}$ is the estimate of $X_{i,j}^{True}$ for simulation v and m is the number of size-classes. Estimation performance for the size–transition matrix was not based on relative errors because some of the entries of the size–transition matrix can be very small and hence the evaluation of estimation performance may be based primarily on the small entries of the size–transition matrix rather than those that constitute the bulk of this matrix. Equation 9 is such that 'perfect' performance (the estimate of the size–transition matrix is identical to the true matrix) would lead to a value of 0, while a complete lack of overlap between the estimated and true matrices would lead to a value of 100.

The performance measure used to summarise the ability to estimate the equilibrium size structure was again the sum of squared absolute errors, i.e. for simulation v:

$$\frac{100}{m} \sum_j |N_j^{True} - N_{j,v}^{Est}| \tag{10}$$

where N_j^{True} is the true proportion of animals in size-class i and $N_{j,v}^{Est}$ is the estimate of N_j^{True} for simulation v.

Performance measures (8)–(10) are such that larger values imply poorer performance. However, they do not indicate the

cause of the poor performance (e.g. is growth estimated to be too fast or too slow?). Therefore, results were reported for a performance measure that quantifies the error associated with the mean size of the equilibrium distribution:

$$\overline{L}^{True}/\overline{L}_v^{True} \tag{11}$$

where $\overline{L}^{True}$ is the mean size of animals at equilibrium:

$$\overline{L}^{True} = \sum_i i N_i^{True} \Big/ \sum_i N_i^{True} \tag{12}$$

and $\overline{L}_v^{Est}$ is the estimate of $\overline{L}^{True}$ for simulation v.

Application to tiger and endeavour prawns in the Gulf of Carpentaria

The data available to estimate the parameters of the size–transition matrix for these prawns are tag–recapture data from experiments conducted in the north-western Gulf of Carpentaria, Australia in 1983 and 1984. These data have been analysed previously for the two tiger prawns (*P. semisulcatus* and *P. esculentus*) by Somers and Kirkwood (1991) and Wang *et al.* (1995), and for blue endeavour prawns, *M. endeavouri*, by Buckworth (1992). In common with Somers and Kirkwood (1991) and Wang *et al.* (1995), the data were restricted to animals that were at liberty for at least 2 weeks and were not infected (at release or recapture) by the bopyrid parasite *Epipenaeon ingens*. Only prawns for which species, sex, length-at-release, length-at-recapture and time-at-liberty were known were included in the analyses.

Results

Simulation analyses

Fig. 2 summarises the distributions for the six performance measures for the eight methods for base-case versions of scenarios #1–#3. In order to identify a set of 'best' methods, the methods were ranked in terms of the values for three of the performance measures, and these ranks were averaged across simulations to determine a single metric of performance (Table 4). According to these ranks, the best method for the base-case simulations is method D, because it achieves the lowest rank in many cases and it is never among the poorest methods. In contrast, methods A,
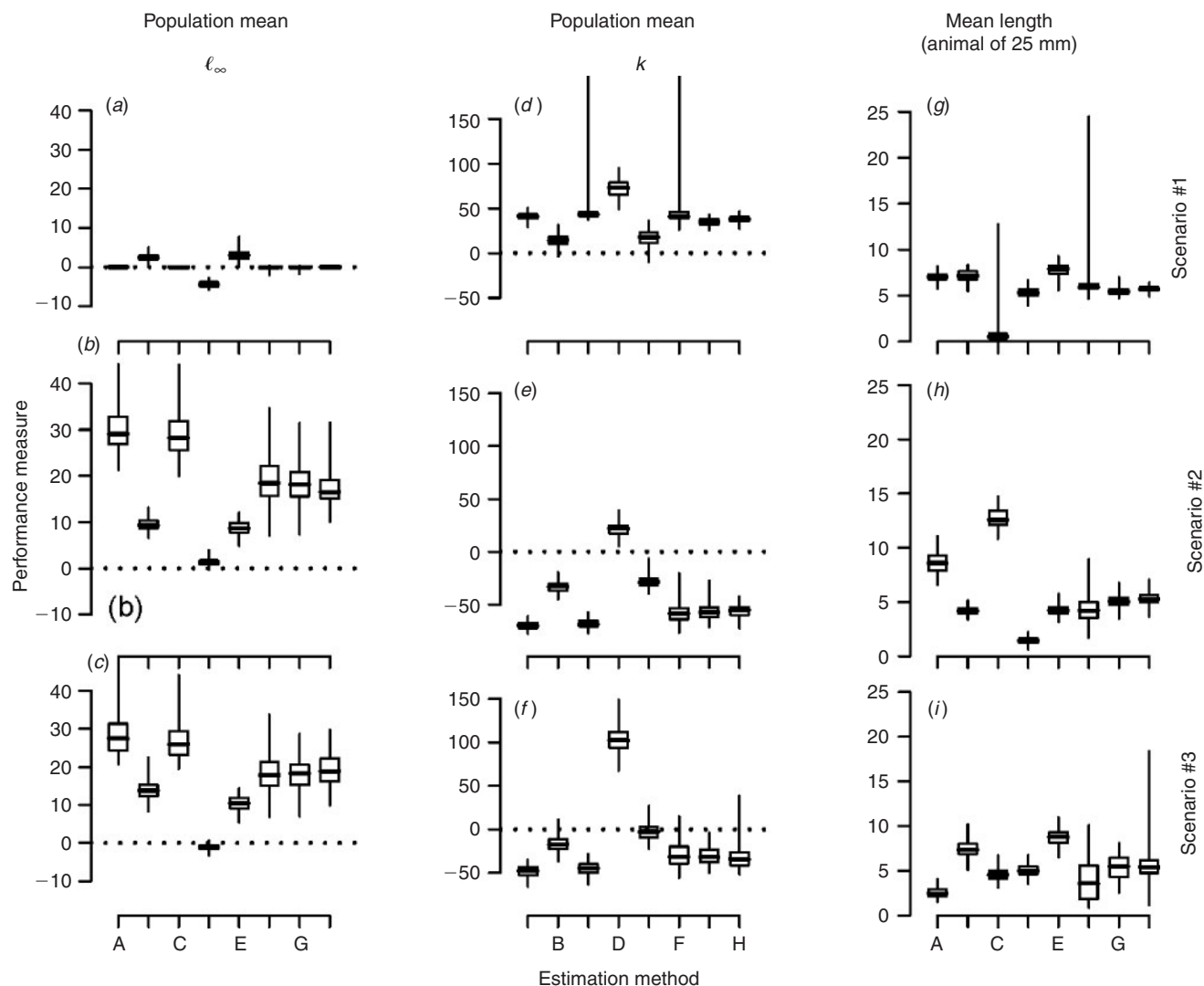
Fig. 2. Distributions of performance measures (medians horizontal line, interquartile range box, range lines) for the population mean value for ℓ_∞, the population mean value for k, the expected size after 10, 20, . . . weeks for animals of 25-mm carapace length (CL), the entries of the size–transition matrix, the equilibrium size distribution, and the population mean length at equilibrium, for eight alternative estimation methods for scenarios #1–#3 and a sample size of 350. Results are shown in (*a*), (*d*), (*g*), (*j*), (*m*) and (*p*) for scenario #1, in (*b*), (*e*), (*h*), (*k*), (*n*) and (*q*) for scenario #2, and in (*c*), (*f*), (*i*), (*l*), (*o*) and (*r*) for scenario #3.

B and C perform best for some scenarios and/or performance measures, but very poorly for others. Methods E, F, G and H are never 'best' for any of the performance measures for the base-case simulations. Method E outperforms methods F, G and H for scenario #2 and method G is almost always better than methods F and H (Table 4).

Further analyses focus on methods C, D and G so as to include the 'best' method for the base-case simulations (D), the method that consistently performed best for scenario #1 (C), and one of the better methods from classes #2 and #3 (method G). Figs 3 and 4 show the distributions for the performance measures for these three methods for the sensitivity tests and Fig. 5 shows the sensitivity of the values for the performance measures for the

base-case version of scenario #3 to the sample size (the results for the other scenarios are qualitatively identical).

Application to data for tiger and endeavour prawns

The applications to the data for tiger and endeavour prawns are based on methods C, D, E and G. Results are shown for method E in Fig. 6 because this method belongs to a different class of method than methods C, D and G. The estimated growth increment over 10 weeks (a period selected so that a change in size is easily detected visually) as a function of length, the estimated standard deviation of growth increment over 10 weeks as function of length, and the equilibrium size-structure of a prawn population for which the rate of total mortality
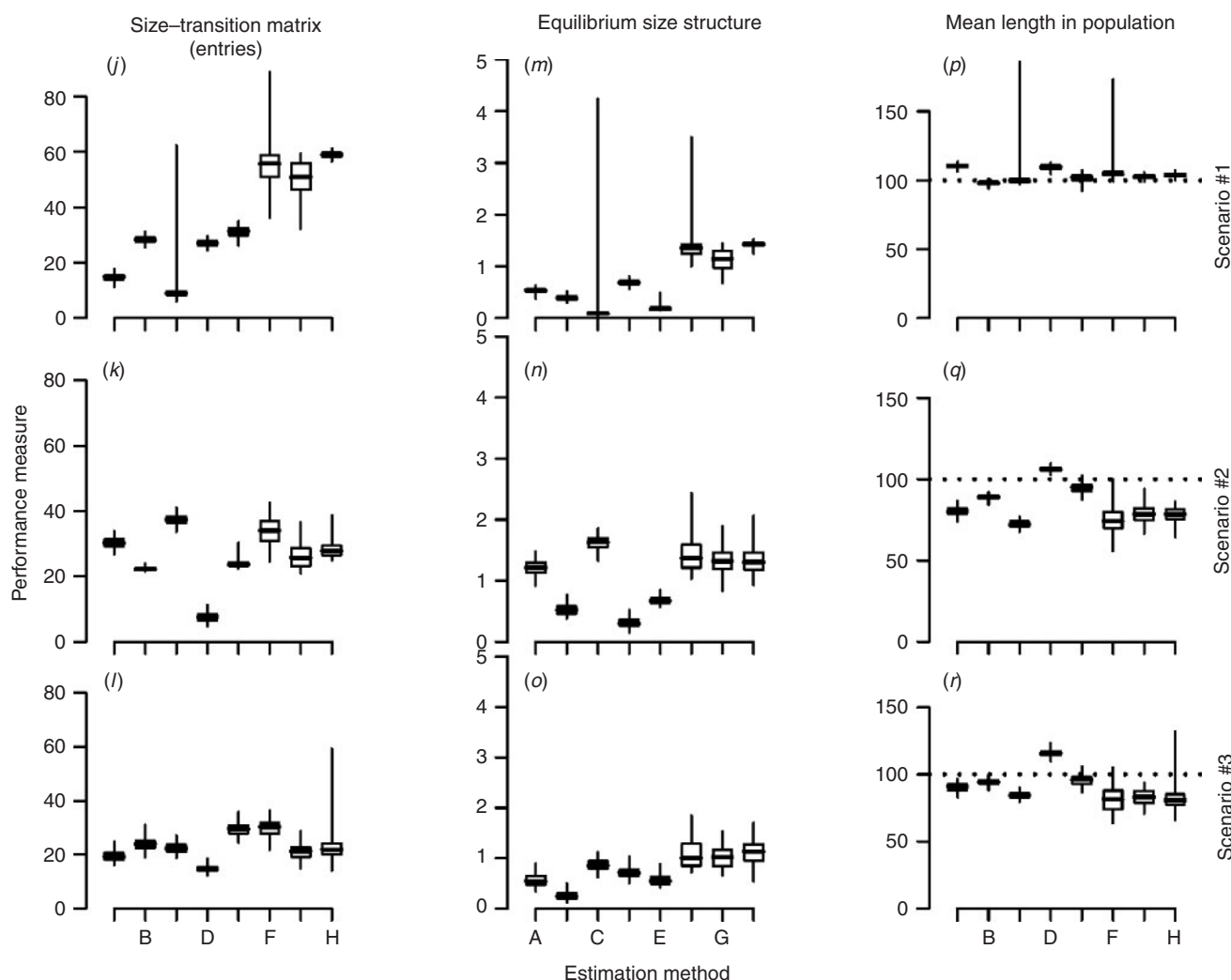
Fig. 2.   (*Continued*)

is 0.045 week^{-1} and for which all recruitment is to the first size-class, differ quite markedly among methods C, D, E and G (Fig. 6). For example, methods D and E tend to indicate that there is an accumulation of density at large size (most evident for *P. semisculcatus* females and *P. esculentus* females) and this behaviour is also evident for method G for both sexes for *P. esculentus* and for *M. endeavouri* females. In contrast, the equilibrium size distribution from method C either declines monotonically with increasing size (all cases except *P. esculentus* males) or is uniform across a broad range of sizes (*P. esculentus* males).

There are differences among methods in the expected growth increment (left panels in Fig. 6) and, particularly, the variation in growth increment (centre panels in Fig. 6). Specifically, method C implies that the greatest variation in growth increment occurs for the small sizes (as expected given the form of Eqn 1 and the assumption of individual variability in k). The standard deviation of growth increment is essentially independent of size

for method D (again expected given the form of Eqn 1 and the assumption of individual variability in ℓ_∞).

The estimated values for k, the age-at-release and ℓ_∞ should conform with the probability distributions underlying Eqns 2a–2d (Wang *et al.* 1995). The estimates of k and ℓ_∞ are well represented by gamma distributions for *P. semisulcatus* females and *P. esculentus* (Fig. 7; the results for the remaining cases are qualitatively identical). The ages-at-release are also well represented by gamma distributions, although that for *P. semisulcatus* females is perhaps not as good as desired (Fig. 7a, lower right panel).

Discussion

Simulation analyses

There is no 'best' method, in the sense that one method consistently leads to the estimates with the best (generally lowest)

	Estimation method							
	A	B	C	D	E	F	G	H
Scenario #1								
Mean length	6.3	6.5	**1.1**	2.7	<u>7.8</u>	4.7	2.7	4.1
Size trans mtrx	2.0	3.9	**1.1**	3.0	5.0	6.9	6.1	<u>8.0</u>
Eqn size struc	4.0	3.0	**1.1**	5.0	2.0	7.0	6.0	<u>7.9</u>
Scenario #2								
Mean length	7.0	3.0	<u>8.0</u>	**1.0**	3.3	3.3	4.9	5.4
Size trans mtrx	5.8	2.2	<u>7.8</u>	**1.0**	3.3	6.8	4.1	5.0
Eqn size struc	4.8	1.9	<u>7.5</u>	**1.1**	3.0	6.3	5.6	5.8
Scenario #3								
Base								
Mean length	**1.4**	6.9	3.6	4.0	<u>7.9</u>	2.8	4.6	4.7
Size trans mtrx	2.6	5.3	4.5	**1.0**	<u>7.3</u>	<u>7.5</u>	3.6	4.2
Eqn size struc	2.8	**1.0**	5.4	3.9	2.8	<u>6.7</u>	6.4	<u>7.1</u>
Prc-1								
Mean length	6.6	<u>7.9</u>	2.1	**1.0**	3.2	6.4	3.7	5.0
Size trans mtrx	3.8	3.0	4.7	5.9	**1.4**	<u>8.0</u>	2.4	7.0
Eqn size struc	6.2	**1.2**	1.9	<u>8.0</u>	5.8	5.4	3.1	4.5
Prc-2								
Mean length	<u>8.0</u>	6.2	6.6	**1.0**	2.6	5.1	2.4	4.1
Size trans mtrx	4.6	**1.6**	3.8	4.6	3.4	<u>8.0</u>	3.0	7.0
Eqn size struc	4.3	**1.0**	2.0	<u>8.0</u>	3.1	6.8	4.6	6.1

values for the performance measures. Rather, different methods perform 'best' for different quantities and scenarios. However, certain patterns do emerge from Fig. 2. For example, it is readily apparent that the ability to estimate the population mean value for ℓ_∞ is generally greater for scenario #1 (individual variability in k) than for scenarios #2 and #3 (individual variability in ℓ_∞, and ℓ_∞ and k respectively), whereas there is noteworthy bias for the population mean value for k for most methods and scenarios. Furthermore, it is also clear that being unable to estimate the population mean values for ℓ_∞ and k adequately does not automatically imply an inability to estimate the elements of the size–transition matrix, primarily because ℓ_∞ and k are negatively correlated. For example, estimation method D is poorest in terms of estimating the population mean values for ℓ_∞ and k for scenario #1 (Fig. 2*a, d*), but outperforms methods E, F, G and H when estimating mean length (Fig. 2*g*), the entries of the size–transition matrix (Fig. 2*j*) and the equilibrium size distribution (Fig. 2*m*) for scenario #1. The results in Fig. 2 also indicate that there is a 'class effect'. For example, the biases of the estimates of ℓ_∞ and k for the three methods based on class #3 (specified standard deviation) always have the same sign as do those for the two methods (A and C) that assume that k varies among individuals.

Method D leads to values for the mean size at equilibrium that are positively biased for all of the base-case simulations (Fig. 2*p*,

q, r), suggesting that it overestimates growth. In contrast, the remaining methods lead to negatively biased estimates of the mean size at equilibrium for scenarios #2 and #3 (Fig. 2*q, r*), which suggests that they underestimate growth rate for this scenario. Method B leads to negatively biased estimates of the mean size at equilibrium for scenario #1, whereas the estimates of this quantity from the remaining methods are positively biased for this scenario (Fig. 2*p*).

Although method D (and C for scenario #1) perform 'best' for base-case simulations in Fig. 2 and Table 4, the assumptions underlying the methods that model individual variation in growth are not identical to those on which the simulated datasets are based. Specifically, in contrast to the assumptions underlying methods C and D, selectivity is a logistic function of length in the simulated population, whereas both k and ℓ_∞ vary among individuals for scenario #3. Estimation performance is not sensitive to the true distribution for individual variation in growth (see the results when k and ℓ_∞ are actually log-normally distributed and when these parameters are normally distributed in Fig. 3; 'Logn' and 'CoCV' respectively). Methods C, D and G generally deteriorate slightly when tagged animals are measured with error (Fig. 3; 'Obs'), although performance is slightly better for some combinations of method and performance measures when there is measurement error. In contrast, allowing for process error in growth (Fig. 3, 'Prc-1' and 'Prc-2') has a large impact on estimation performance and the relative performance of the various methods changes (Fig. 3; Table 4). Specifically, method D performs poorest of the eight methods in terms of estimating the equilibrium size distribution and is no longer best at estimating the entries of the size–transition matrix when there is process error in growth (Table 4).

The ability to estimate the entries of the size–transition matrix depends on the age-structure of the tagged prawns. For example, estimation performance deteriorates markedly (a tripling in the value of Eqn 9 in some cases) if all the tagged animals were 20 weeks old (Fig. 4; 'All20'). This reflects, in part, a reduction in the ranges of tagged animals in the dataset, but also that the larger tagged animals are now the faster-growing individuals, thereby violating the assumption for methods A–D that the distributions for ℓ_∞ and k are independent of size-class. Performance is generally better (lower values for the performance measures in 12 of 18 cases in Fig. 4) when selectivity is uniform. The performance of method D improves noticeably when selectivity is uniform for scenario #2 because, in this case, the estimation method matches how the data are generated. Method D is able to estimate the population mean values for ℓ_∞ and k with almost no bias for this case (results not shown).

Although the widths of the distributions drop with increasing sample size, there is only strong evidence for improved estimation performance for method G when the entries of the size–transition matrix are estimated (Fig. 5*f*). The patterns evident in Figs 3–5 therefore appear to be a consequence of the structural differences between the assumptions on which each method is based and the way the pseudo datasets were generated, in particular how individual variation in growth was represented, the relationship between the age-at-release and the parameter considered to vary among individuals, and whether selectivity is uniform.
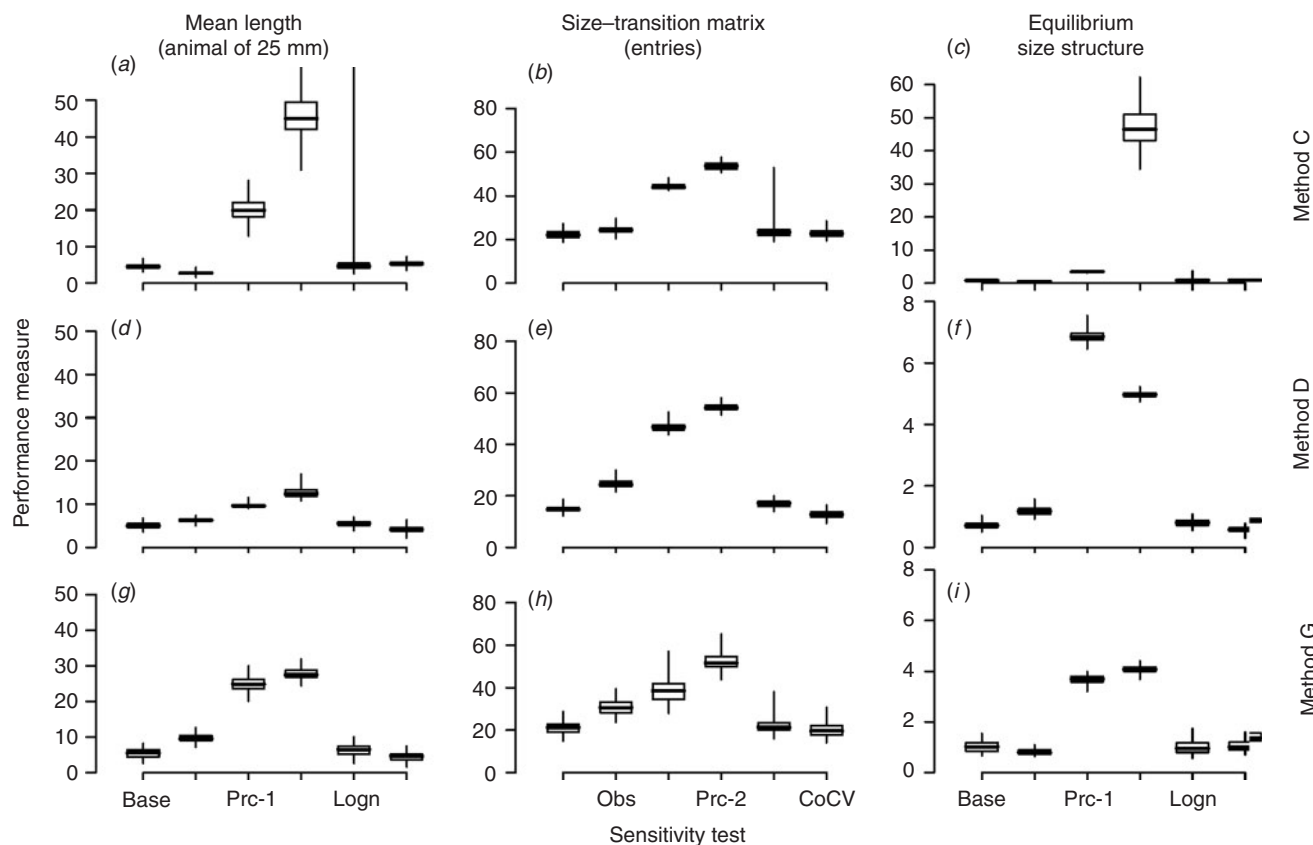
Fig. 3. Distributions of performance measures for the expected size after 10, 20, . . . weeks for animals of 25-mm carapace length (CL), the entries of the size–transition matrix and the equilibrium size distribution for the base-case version of scenario #3 and five variants thereof (see Table 3). Results are shown in (*a*, *b*, *c*) for method C, in (*d*, *e*, *f*) for method D, and in (*g*, *h*, *i*) for method G. The sample size is 350 for all trials.

Application to data for tiger and endeavour prawns

The estimates for the age-at-release from estimation method C seem unrealistically high (>>50 weeks) for a prawn (Fig. 7). (Many of the *P. semisulcatus* individuals tagged during the experiment were very large and could conceivably be age 1+, but the ages implied by method C are implausible for a prawn.) This can be attributed to the nature of the estimator when k is assumed to vary among individuals. Specifically, estimation method C implicitly assumes that ℓ_∞ is larger than the size of the largest recaptured prawn (e.g. 53 mm for female *P. semisulcatus*), which implies that k must be fairly low and hence that the age-at-release must be high. The possibility of estimation method C leading to positively biased estimates of ℓ_∞ and negatively biased estimates of k in the face of individual variation in ℓ_∞ was identified in the simulations (Fig. 2*b*, *e*). In contrast, the impact of individual variation in k on the estimates from method D is much smaller (Fig. 2*a*, *d*).

The estimates of the population mean values for ℓ_∞ and k for *P. semisulcatus* from method D (36.5 mm, 0.076 week^{-1} for males; 47.5 mm, 0.059 year^{-1} for females) are similar to those obtained by Wang *et al.* (1995) (36 mm, 0.081 week^{-1} for males; 46.6 mm, 0.060 year^{-1} for females), with the differences

attributable to slight difference in dataset choice. The differences between the estimates from the present study and those from Somers and Kirkwood (1991) (37.5 mm, 0.062 week^{-1} for males; 51.5 mm, 0.043 year^{-1} for females), particularly those for k, are more substantial. Somers and Kirkwood (1991) estimated growth parameters using an approach in which ℓ_∞ was assumed to be distributed according to a truncated normal distribution, but did not account for individual variation in age-at-release. The approach of Somers and Kirkwood (1991) is therefore most similar to method B, the estimates from which for *P. semisulcatus* are 37.5 mm, 0.058 week^{-1} for males and 50.6 mm, 0.042 year^{-1} for females.

General discussion

There are a large number of ways of constructing size–transition matrices for use in size-structured stock assessment methods. The results of the present paper indicate that the relative performance (and robustness) of these methods is not equal. It is well known that estimating the parameters of a von Bertalanffy growth curve ignoring individual variation in growth leads to biased estimates of the population mean values for k and ℓ_∞
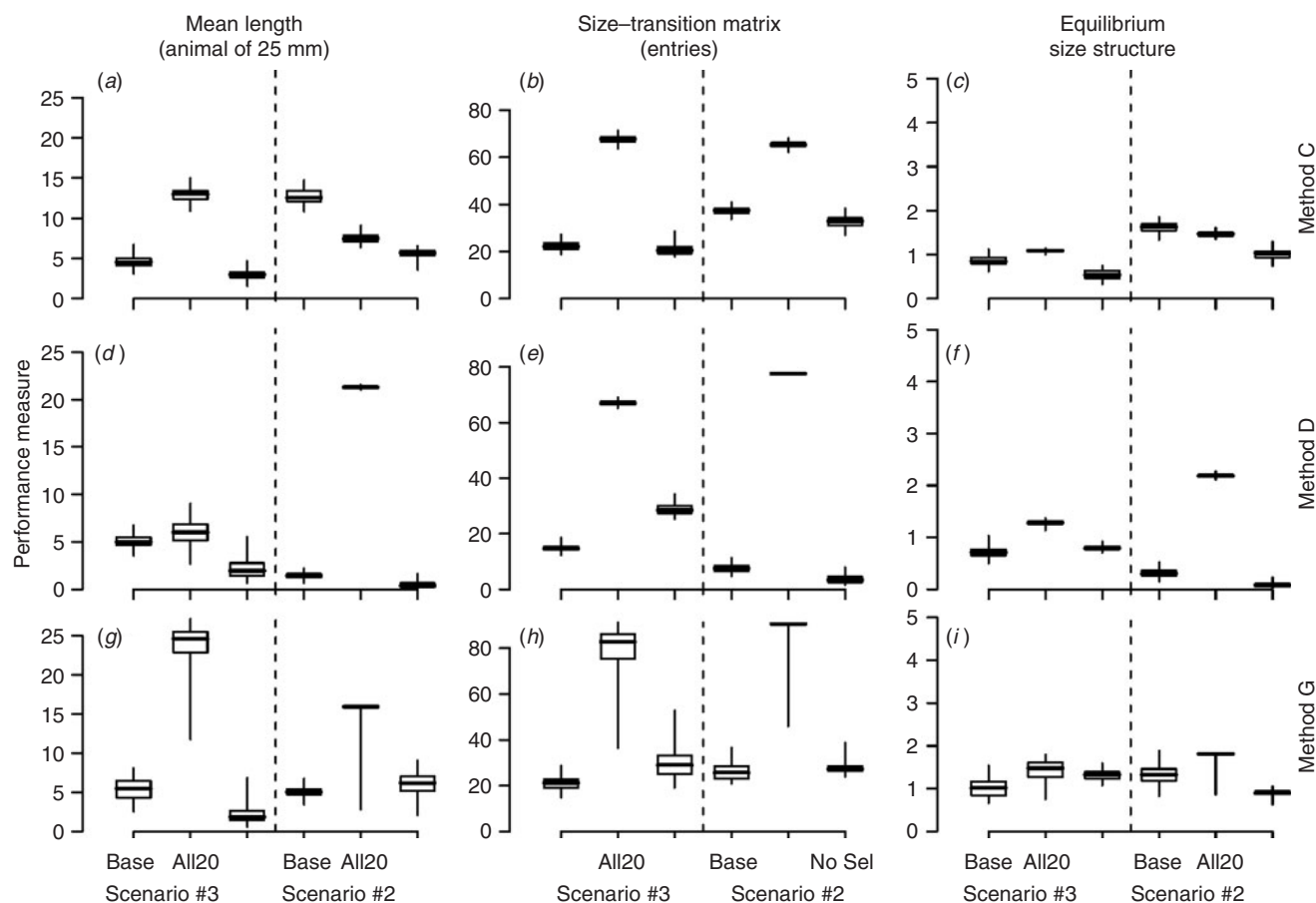
Fig. 4. Distributions of performance measures for the expected size after 10, 20, . . . weeks for animals of 25-mm carapace length (CL), the entries of the size–transition matrix and the equilibrium size distribution for three variants of scenarios #2 and #3 (see Table 3 for abbreviation definitions) and a sample size of 350. Results are shown in (*a*, *b*, *c*) for method C, in (*d*, *e*, *f*) for method D, and in (*g*, *h*, *i*) for method G.

(e.g. Sainsbury 1980). Methods such as F, G and H are essentially based on such methods and, consistent with expectations, lead to severely biased estimates of both population mean values as well as of the entries of the size–transition matrix. Methods that allow for individual variation in either k or ℓ_∞ perform best, particularly when account is also taken of variation in the age-at-release. This conclusion will depend to some extent on the scenarios underlying the simulations, although the scenarios considered in the present paper were designed to mimic reality to the maximum extent possible by examining scenarios in which k or ℓ_∞ (or both) vary among individuals, allowing for population dynamic process errors such as natural mortality and variability in dates of spawning and by allowing for selection during both tagging and recapture.

Of the four methods that allow for individual variation in growth, that which assumes that ℓ_∞ varies among individuals and accounts for variability in the age-at-release (method D) is the most robust (in particular, it is fairly robust to the assumption that it is ℓ_∞ rather than k that varies among individuals; Fig. 2). In contrast, assuming that k varies among individuals when, in fact, ℓ_∞ or both k and ℓ_∞ do, can lead to severely biased estimates

of the entries in the size–transition matrix (Fig. 2) as well as unrealistic estimates of the distribution of age-at-release (Fig. 7).

The performance of method D is poorer than indicated by Wang *et al.* (1995). This is attributable to several causes: (1) allowance is made in the simulations of this paper for selectivity, and (2) the aim of the methods in this paper is to estimate the parameters of the size–transition matrix rather than k and the population mean value of ℓ_∞. The latter is important because fewer assumptions are needed to estimate the population mean values for the growth parameters than the entries of the size–transition matrix.

The simulation study of the present paper did not explore the possibility of seasonal variation in k, that the growth curve changes at, for example, maturity, or the possibility that the growth rates for tagged animals differ from those for untagged animals. The first two factors relate to model uncertainty. In principle, any of the methods outlined in the present paper could be implemented based on alternative relationships between growth increment and current size, such as the general growth curve of Schnute (1981), growth curves that include seasonal growth (e.g. Pauly 1990; Pauly *et al.* 1992) and growth
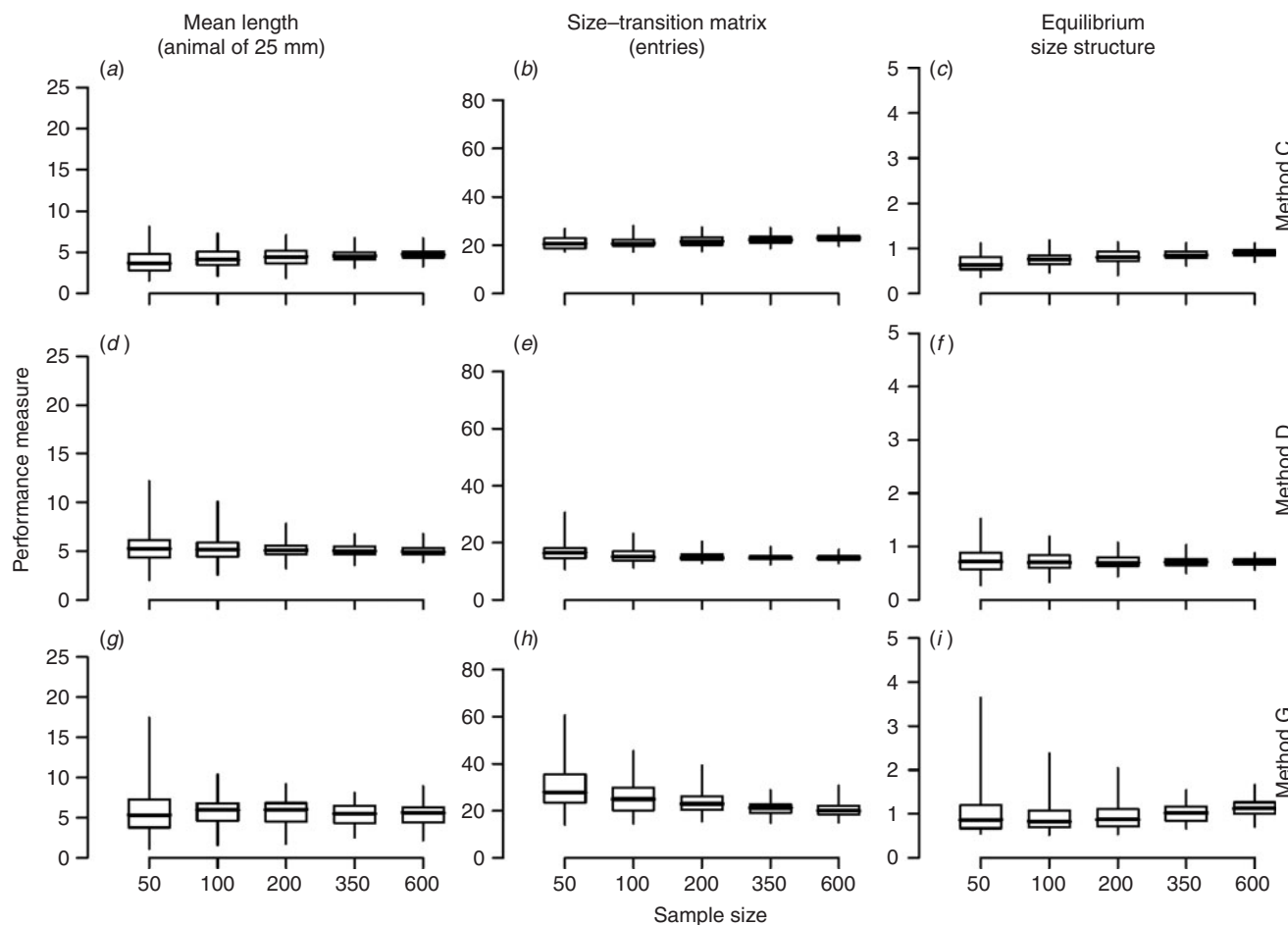
Fig. 5. Distributions of performance measures for the expected size after 10, 20, ... weeks for animals of 25-mm carapace length (CL), the entries of the size–transition matrix, and the equilibrium size distribution, for scenario #3 for different sample sizes. Results are shown in (*a*, *b*, *c*) for method C, in (*d*, *e*, *f*) for method D, and in (*g*, *h*, *i*) for method G.

curves that exhibit a change in growth, for example at maturity (e.g. McGarvey *et al*. 1999). Model selection methods (such as AIC or BIC) could then be used to select among the alternatives. Naturally, additional simulation work would be needed to assess the power to select among alternative growth models. The impact of tagging on growth rate can be an important issue, but would impact all of the methods considered in this paper. The simulations are all based on the assumption of continuous growth. Future simulation studies could include pseudo datasets generated using models that explicitly allow for moulting.

Further extensions to this work are possible. In particular, consideration could be given to a method in which both k and ℓ_∞ vary among individuals (although the difference in results between scenarios #2 and #3, at least for estimation method D, suggests that the benefits of developing such a method may be limited compared with its computational cost; see also the results obtained by Eveson *et al*. (2007)).

An additional further development would be to use model-selection methods (e.g. AIC or BIC) to select among alternative assumptions regarding which parameters should be considered to vary among individuals and the probability distribution(s) for these parameters, although the results obtained by Eveson *et al*. (2007) suggest that these model selection techniques may not perform very well. Finally, consideration could be given to moving to a Bayesian estimation framework to enable *a priori* information (such that ages of tagged prawns are not likely to be >50 weeks) to be explicitly included in the analyses.

Finally, no attempt has been made in this paper to integrate estimation of the parameters of the size–transition matrix directly into a stock assessment. While this is straightforward for approaches based on classes 2 and 3, this is not the case for the methods that assume that either k or ℓ_∞ varies among individuals because of the need to compute the entries in the size–transition matrix using numerical methods.
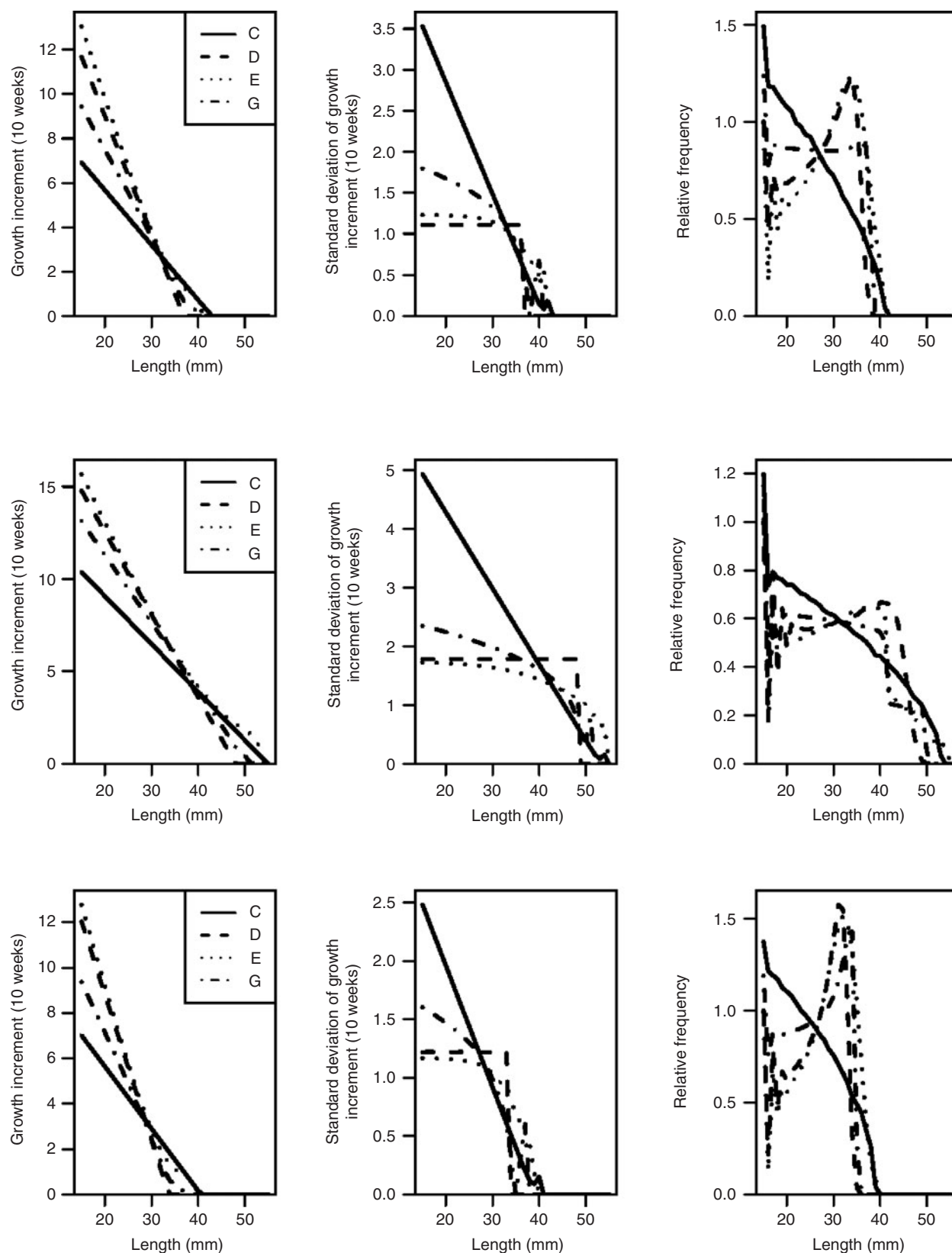
Fig. 6. Expected growth increment (mm) and the standard deviation of growth increment over 10 weeks (mm) for three prawn species, and the implied equilibrium size compositions when total mortality $= 0.045$ week^{-1} for four of the methods for estimating the size–transition matrix (C, D, E and G). Results are shown for males in the upper panels and for females in the lower panels.
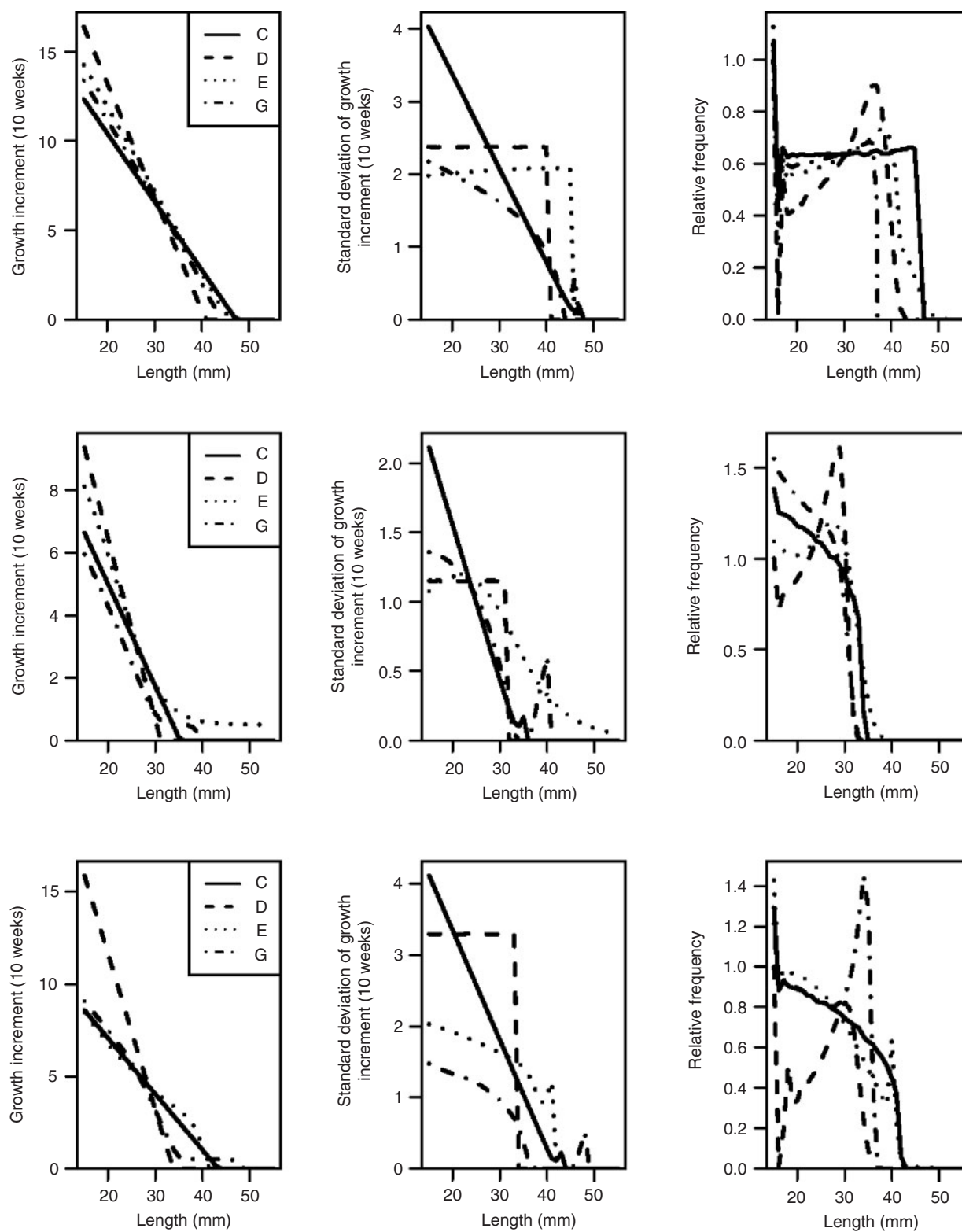
Fig. 6. (*Continued*)
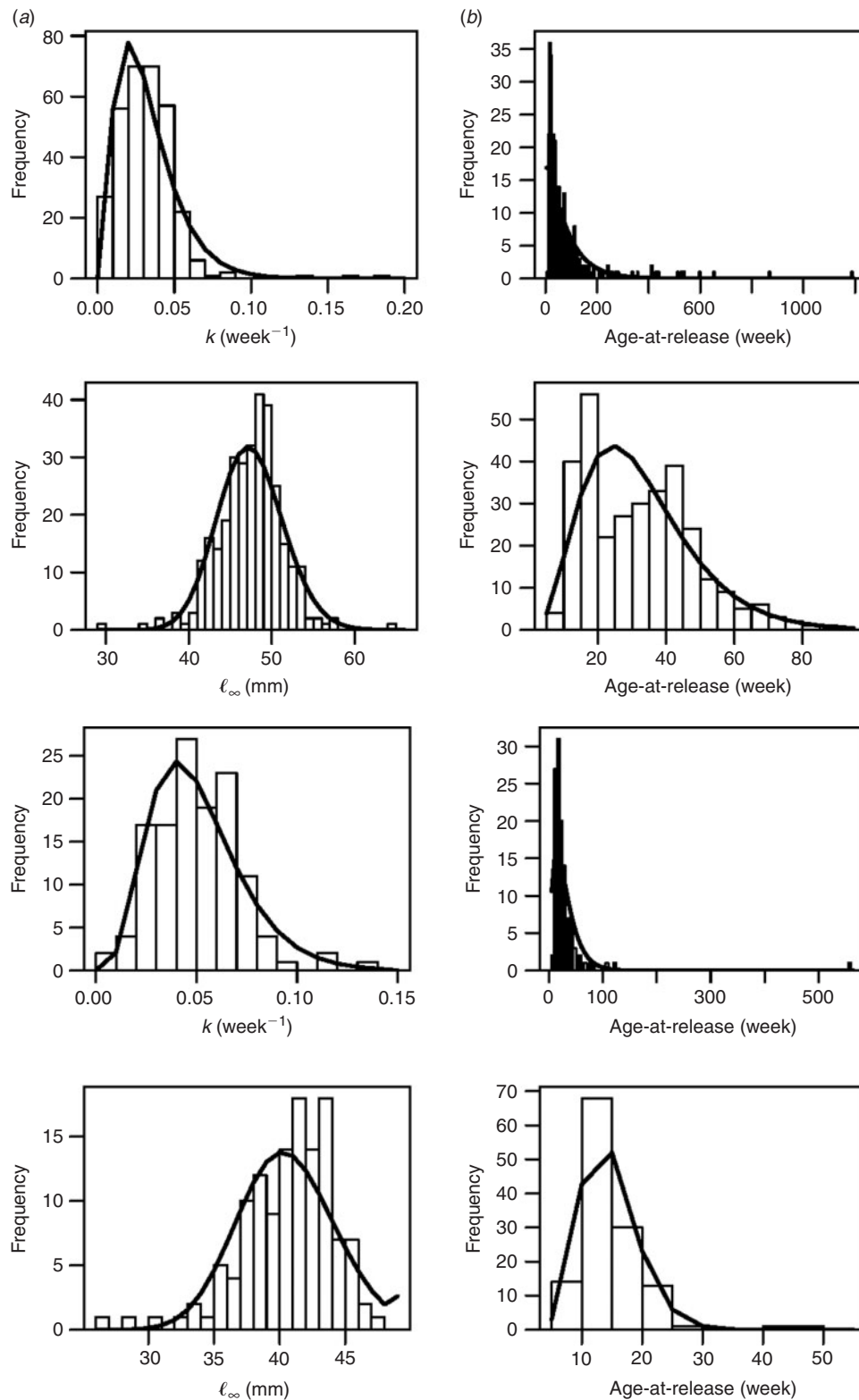
Fig. 7. Estimated values for k and the age-at-release (estimation method C) and ℓ_∞ and the age-at-release (estimation method D) (upper and lower panels for each species) (bars) along with the estimated probability density functions based on the estimated values for the parameters of the von Bertalanffy growth curve (lines). Results are shown in (*a*) for female *Penaeus semisulcatus* and in (*b*) for female *Penaeus esculentus*.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**, 716–723. doi:10.1109/TAC.1974.1100705

Brandão, A., Butterworth, D. S., Johnston, S. J., and Glazer, J. P. (2004). Using a GLMM to estimate the somatic growth rate trend for west coast rock lobster. *Fisheries Research* **70**, 339–349. doi:10.1016/J.FISHRES.2004.08.012

Breen, P. A., Kim, S. W., and Andrew, N. L. (2003). A length-based Bayesian stock assessment model for the New Zealand abalone *Haliotis iris*. *Marine and Freshwater Research* **54**, 619–634. doi:10.1071/MF02174

Buckworth, R. C. (1992). Movements and growth of tagged blue endeavour prawns *Metapenaeus endeavouri* (Schmitt 1926) in the western Gulf of Carpentaria, Australia. *Australian Journal of Marine and Freshwater Research* **43**, 1283–1299. doi:10.1071/MF9921283

Burnham, K. P., and Anderson, D. R. (2002). 'Model Selection and Inference: A Practical Information-Theoretic Approach.' 2nd edn. (Springer-Verlag: New York.)

Chen, Y., Kanaiwa, M., and Wilson, C. (2005). Developing and evaluating a size-structured stock assessment model for the American lobster, *Homarus americanus*, fishery. *New Zealand Journal of Marine and Freshwater Research* **39**, 645–660.

de Lestang, S., and Melville-Smith, R. (2006). West coast rock lobster managed fishery status report. In 'State of the Fisheries Report 2005/06'. (Eds W. J. Fetcher and F. Head.) pp. 14–21. (Western Australian Department of Fisheries: Perth.)

Dichmont, C. M., Punt, A. E., Deng, A., Dell, Q., and Venables, W. (2003). Application of a weekly delay-difference model to commercial catch and effort data in Australia's Northern Prawn Fishery. *Fisheries Research* **65**, 335–350. doi:10.1016/J.FISHRES.2003.09.024

Eveson, J. P., Polacheck, T., and Laslett, G. M. (2007). Consequences of assuming an incorrect error structure in von Bertalanffy growth models: a simulation study. *Canadian Journal of Fisheries and Aquatic Sciences* **64**, 602–617. doi:10.1139/F07-036

Food and Agriculture Organization (2008). Statistics. Available online at http://www.fao.org/corp/statistics/en/ [Verified January 2008].

Galeano, D., Vieira, S., Shafron, W., and Newton, P. (2006). Australian fisheries surveys report 2005. ABARE report prepared for the Fisheries Research Fund, Canberra.

Hobday, D., and Punt, A. E. (2001). Size-structured population modelling and risk assessment of the Victorian southern rock lobster, *Jasus edwardsii*, fishery. *Marine and Freshwater Research* **52**, 1495–1507. doi:10.1071/MF01050

Kim, S. W., Bentley, N., Starr, P. J., and Breen, P. A. (2004). Assessment of red rock lobsters (*Jasus edwardsii*) in CRA 4 and CRA 5 in 2003. New Zealand Fisheries Assessment Report 2004(8).

Laslett, G. M., Eveson, J. P., and Polacheck, T. (2002). A flexible maximum likelihood approach for fitting growth curves to tag–recapture data. *Canadian Journal of Fisheries and Aquatic Sciences* **59**, 976–986. doi:10.1139/F02-069

Maunder, M. N. (2001). Growth of skipjack tuna (*Katsuwonus pelamis*) in the eastern Pacific Ocean, as estimated from tagging data. *Bulletin of the Inter-American Tropical Tuna Commission* **22**, 95–131.

McGarvey, R., and Feenstra, J. E. (2001). Estimating length-transition probabilities as polynomial functions of premoult length. *Marine and Freshwater Research* **52**, 1517–1526. doi:10.1071/MF01172

McGarvey, R., Ferguson, G. J., and Prescot, J. H. (1999). Spatial variation in mean growth rates at size of southern rock lobster, *Jasus edwardsii*, in South Australian waters. *Marine and Freshwater Research* **50**, 333–342. doi:10.1071/MF97172

McLoughlin, K. (2006). Northern Prawn Fishery. In 'Fishery Status Reports 2005: Status of Fish Stocks Managed by the Australian Government'. (Ed. K. McLoughlin.) pp. 23–34. (Bureau of Rural Sciences: Canberra.)

Pauly, D. (1990). Length converted catch curves and the seasonal growth of fishes. *ICLARM Fishbyte* **8**, 33–38.

Pauly, D., Soriano-Bertz, M., Moreau, J., and Jarre-Teichmann, A. (1992). A new model accounting for seasonal growth cessation in fishes. *Australian Journal of Marine and Freshwater Research* **43**, 1151–1156. doi:10.1071/MF9921151

Punt, A. E., Kennedy, R. B., and Frusher, S. (1997). Estimating the size-transition matrix for Tasmanian rock lobster, *Jasus edwardsii*. *Marine and Freshwater Research* **48**, 981–992. doi:10.1071/MF97017

Punt, A. E., Hobday, D., Gerhard, J., and Troynikov, V. S. (2006). Modelling growth of rock lobsters, *Jasus edwardsii*, off Victoria, Australia using models that allow for individual variation in growth parameters. *Fisheries Research* **82**, 119–130. doi:10.1016/J.FISHRES.2006.08.003

Sainsbury, K. J. (1980). Effect of individual variability on the von Bertalanffy growth equation. *Canadian Journal of Fisheries and Aquatic Sciences* **37**, 241–247. doi:10.1139/F80-031

Schnute, J. (1981). A versatile growth model with statistically stable parameters. *Canadian Journal of Fisheries and Aquatic Sciences* **38**, 1128–1140. doi:10.1139/F81-153

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics* **6**, 461–464. doi:10.1214/AOS/1176344136

Smith, M. T., and Addison, J. T. (2003). Methods for stock assessment of crustacean fisheries. *Fisheries Research* **65**, 231–256. doi:10.1016/J.FISHRES.2003.09.017

Somers, I. F., and Kirkwood, G. P. (1991). Population ecology of the grooved tiger prawn, *Penaeus semisulcatus*, in the north-western Gulf of Carpentaria, Australia: growth, movement, age structure and infestation by the bopyrid parasite *Epipenaeon ingens*. *Australian Journal of Marine and Freshwater Research* **42**, 349–367. doi:10.1071/MF9910349

Sullivan, P. J., Lai, H. L., and Gallucci, V. F. (1990). A catch-at-length analysis that incorporates a stochastic model of growth. *Canadian Journal of Fisheries and Aquatic Sciences* **47**, 184–198. doi:10.1139/F90-021

Troynikov, V. S. (1998). Probability density functions useful for parameterization of heterogeneity in growth and allometry data. *Bulletin of Mathematical Biology* **60**, 1099–1121. doi:10.1006/BULM.1998.0058

Wang, Y.-G., Thomas, M. R., and Somers, I. F. (1995). A maximum likelihood approach for estimating growth from tag-recapture data. *Canadian Journal of Fisheries and Aquatic Sciences* **52**, 252–259. doi:10.1139/F95-025